PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	History of coronary heart disease increased the mortality rate of COVID-19 patients: a nested case-control study
AUTHORS	Gu, Tian; Chu, Qiao; Zhang-Sheng, Yu; Fa, Botao; Li, Anqi; Xu, Lei; Wu, Ruijun; He, Yaping

VERSION 1 – REVIEW

REVIEWER	Chien-Cheng Huang Chi Mei Medical Center, Taiwan
REVIEW RETURNED	19-Apr-2020

GENERAL COMMENTS	The authors conducted a nested case-control study to evaluate the risk of pre-existing comorbidities on COVID-19 mortality. In total, 94 cases and 181 controls were recruited into this study. They found that history of comorbidities significantly increased the death risk of COVID-19. The estimated mortality risk in patients with CHD was three times of those without CHD. Older age was also associated with increased death risk: every 5-year increase in age was associated with a 20% increased risk of mortality (p<0.001). They concluded that extra care and early medical intervention are needed for patients with pre-existing comorbidities, especially CHD. In general, this study was well conducted and the manuscript was well written. I have the following comments. Abstract. Please describe the distribution of age and gender in both groups briefly in the Results. Table 1. There were three columns of HRs for multivariate model. Please describe what variable each column was adjusted for. Discussion. The authors did not recruit treatment into the analyses, which may confound the results. Please explain this issue in the Discussion.

REVIEWER	Priscila Maria Stolses Bergamo Francisco State University of Campinas São Paulo/Brazil
REVIEW RETURNED	19-Apr-2020

GENERAL COMMENTS	**General comments** *The study addresses a relevant question to Public Health today and assesses the relationship between pre-existing comorbidities and mortality from COVID-19, through a nested case-control design and public data from patients with confirmed SARS-CoV infection -2 in*

mainland China outside of Hubei Province. Secondary public domain data were used and the data collection, for each case, was properly described in stages. The model's assumptions were verified to perform statistical analyzes. The method is clearly described and the discussion is well structured.

In Introduction, the first paragraph, the authors could update the number of cases and countries affected around the world, as well as in the 2nd paragraph, including the reference (WHO).

Especific comments

Model Results

Paragraph 2
"Table 2 presents the results of univariate and multivariate weighted Cox models. Older age was associated with significantly higher death risk with similar magnitude in univariate and multivariate models. In the adjusted model, every 5-year increase in age was associated with na estimated 20% higher risk of death (p<0.001). No significant hazard difference was found between male and female patients. Disease infection during the early no-intervention period was associated with a higher risk of death, although not statistically significant".

I suggest:
"Table 2 presents the results of univariate and multivariate weighted Cox models. Older age was associated with significantly higher death risk with similar magnitude in univariate and multivariate models. No significant hazard difference was found between male and female patients. **Although the disease infection during the early no-intervention period was associated with a higher risk of death (pontual estimate), it was not statistically significant."**

"In the adjusted model, every 5-year increase in age was associated with an estimated 20% higher risk of death" - **should not be related with Table 2, since this result is not represented there.**

Paragraph 3

"The increasing number of cumulative comorbities as associated with higher mortality risk in both unadjusted and adjusted models (p<0.001) - In Table 2 and Table S2 the authors present the variable *"Total numbers of comorbities"*, but does not show that the cumulative number of comorbities is related to higher mortality, since this variable does not represent an increasing number of cumulative comorbities (gradient) - **I suggest rewriting the sentence or showing these results inTable S2.**

Paragraph 3

<table>
<tr><td></td><td>All preexisting comorbidities were associated with a higher risk of COVID-19 mortality in the univariate model, of which CHD had the largest hazard ratio (HR) of 4.2 (p<0.001), followed by cerebral infarction (HR=2.9, p=0.004), COPD (HR=2.6, p=0.01), renal failure (HR=2.3, p=0.09), cardiac failure (HR=1.9, p=0.1), history of surgery (HR=1.7, p=0.34), hypertension (HR=1.4, p=0.17), diabetes (HR=1.1, p=0.61) and chronic bronchitis (HR=1.1, p=0.55), **but not all statistically significant**.</td></tr>
</table>

REVIEWER	Moritz Tobiasch Landeskrankenhaus Hall in Tirol, Department of Medicine, Tyrolia, Austria
REVIEW RETURNED	04-May-2020

GENERAL COMMENTS	Thank you very much for the opportunity to review this paper. From a methodological point of view, it is quite obvious (and very welcome) that a fair amount of work and consultation with experts in biostatistics was put in this project.

As I was asked to primarily focus on statistical questions, I would like to specifically address this field - however, there is one more general objection that need to be clarified all along: The process of data gathering via news outlets and websites is all but unconventional for people accustomed to EU data protection standards. In my opinion, it would be advisable to clarify whether there was supervision by any sort of ethics committee.

As far as statistics are concerned, I would like to ask, with all due respect, for some clarifications:
- In the methods section, it is mentioned that for continuous variables, a t-test was used. This test has gone a little out of fashion as it relies on assumptions that are only rarely met, namely, normal distribution and equal variances in both groups. Have all continuous variables been tested for normality, e.g., by Komolgorow-Smirnow or Shapiro-Wilk tests? If the baseline assumptions of normality and equal variances are not met, I would advise to use a non-parametric test.
- I am quite impressed to see such an analysis performed in R. R is more and more used in a dedicated integrated development environment (IDE) such as RStudio. If this was the case, I would like to ask to state this as well. Furthermore, one package was explicitly mentioned and cited, which I greatly appreciate. If other packages were used as well, I would like to cite them as well.
- Is it possible to clarify how preexisting conditions were defined, e.g., for heart failure, was there a LVEF cutoff, or, for hepatic failure, was a classification according to AASLD or EASL or APASL used - or did all these data solely rely on the discretion of treating physicians or the press? The latter would no be a particular worrisome challenge, it just needs to be stated.
- Demographics: Can you comment on the fact why it happened that even after matching, the control cohort proved to be significantly younger than the case cohort?
- The stepwise risk stratification for age with increments of 5 years in a cohort with a median age of 70.7 is a little unprecise, in my view. In my country, the average age of death in males is currently 82 years, which makes up for only a little more than two incremental steps in your model, and I would think that even without COVID-19, between the ages of 70.7 and 82, the death rate might rise by 20% every 5 years.

- Your statement of equal risks of death between male and female patients is interesting and needs some discussion, as per usual assumption in the public (and hereby I mean specialists in critical care medicine treating COVID-19 patients) males are more often dying from COVID-19 than women. Is there, to your knowledge, data available reproducing your finding?
- The last and possibly most important point is still a statistical one, but it happens to affect basically all studies on acute critical conditions. Your model of a 65 year old female with CHD and COVID-19 comes up with quite harsh prognostics quo ad vitam, with a median (!) survival of mere 34 days. From my humble experience, I would rather say that this might be a selection or matching artifact (by the way, the "German anomaly" with exceptionally low mortality despite equal infection numbers in GER compared to other EU countries like ITA, ESP, FRA is possibly quite the inverse effect). May I ask to go through your matching process to find out why exactly such a short survival time was found? What is the time course and prognosis in oligosymptomatic COVID-19 patients with CHD?

Additionnally, I would like to see at least one typo addressed:
- p12 line 14: "plaque" instead of "plague"

As far as references are concerned, in the last days epidemiologic evidence from EU countries, UK, and USA emerged. For an even better discussion, I would like to ask whether the work presented here could be put in some context. Furthermore, a current hypothesis in COVID-19 induced critical illness is a generalised affection of endothelia, which makes preexisting vascular conditions an at least hypothetically dangerous influencing factor. Maybe the discussion could be expanded to these recent findings.

In total, the work is very balanced, thorough, and interesting. After addressing the above mentioned points, the paper might be ready for publication.

REVIEWER	Luis Puente-Maestu Hospital Universitario Gregorio Marañón-Facultad de Medicina de la Universidad Complutense de Madrid
REVIEW RETURNED	09-May-2020

GENERAL COMMENTS	1) The study has some flaws. The most important is the large number of individuals with incomplete data and the important number of excluded cases for the final analysis, since it questions the validity of the study and it seems to this reviewer difficult to solve. 2) Even in waivered from informed consent, the study should have been reviewed by an ethics in human research committee. 3) Page 6, line 32: "The data collection procedure was blinded to patient comorbidity information". Clarify; how the data collection could be blinded to comorbidities if they were considered exposition factors 4) Page 6, line 40: was matched with up to three controls on gender and age ± 1 year old (94 cases, 181 controls). The authors should explain bit more in depth how the sample was collected, randomly among all the possible controls fulfilling the matching criteria, stratified by regions. Furthermore from figure 1 it seem that most of the available control population (i.e 12555/12861) had incomplete

records. If that is true the validity (representativeness) of the study sample would be largely questionable

5) Page 6,line 46: We routinely searched for daily news and public health reports on confirmed COVID-19. Are the final sources of information available for other investigators to check your results?

6) Page 7, line 8: "who had complete information on basic demographics (age, gender and region), disease onset date--the first time a patient became symptomatic, and history of comorbidities (include but not limited to hypertension, cardiovascular disease, diabetes and respiratory diseases) were included in the analysis" Was the proportion of excluded subjects even among regions, or said in other words, had the registry the same quality among regions?. Had all the included individuals the same chance of being diagnosed of some of the comorbidities with independence that the cases came for urban and rural areas, more or less developed zones?

7) Page 7, line 13 "include but not limited to hypertension, cardiovascular disease, diabetes and respiratory diseases) were included in the analysis" all the comorbidities included in the analysis should be shown here or in the supplementary material. AS per table 2 they are not many more.

8) Page 7, line 13: Asymptomatic patients were not included in this study. This must be moved to the study design section. A reference to this criteria has to be included in the abstract and somehow if possible in the title.

9) Page 8, line 6 The matching between cases and controls, and relative weights were simultaneously obtained via KMprob function in multipleNCC R package" The tiem-fixed IP exposure weights should be presented somewhere

10) Page 8,line 10: Asymptomatic patients were not included in this study

11) Page 8,line 12: Those survivors with sampling probabilities of zero were considered as "fail to match" and excluded from the study. The number of excluded patients and the region they come from should be shown.

12) Page 8, line 15 The total number of comorbidities was defined as the summation of comorbidities, ranging from zero to four or above." Was this resulting variable coded for multivariable analysis (i.e $X_i=1$, the rest 0, resulting in k-1 coded variables). Besides it seems from your results (see below), that the number of comorbidities is not an homogeneous risk factor. You may approach to comorbidity analysis by calculating the Carlson index which tries to account for the difference impact of different comorbidities in survival.

13) Table 2 Some HR are shown as 1.0 (1.0-1.1), please give the number of decimals needed to show that both the HR and the confidence interval do not include 1)

14) Pag 11, line 22 The increasing number of cumulative comorbidities was associated with higher mortality risk in both unadjusted and adjusted models ($p<0.001$). This statement may not be true. According to your data does not seem that it is the same to have two comorbidities that were not CHD, that only one which was CHD.

15) Page 11, line 41 "All preexisting comorbidities were associated with a higher risk of COVID-19 mortality in the univariate model," As far as table 2 reports this is not true. Some commodities were not found significantly associated to survival expectancy

Reviewer: 1
Reviewer Name: Chien-Cheng Huang
Institution and Country: Chi Mei Medical Center, Taiwan

The authors conducted a nested case-control study to evaluate the risk of pre-existing comorbidities on COVID-19 mortality. In total, 94 cases and 181 controls were recruited into this study. They found that history of comorbidities significantly increased the death risk of COVID-19. The estimated mortality risk in patients with CHD was three times of those without CHD. Older age was also associated with increased death risk: every 5-year increase in age was associated with a 20% increased risk of mortality (p<0.001). They concluded that extra care and early medical intervention are needed for patients with pre-existing comorbidities, especially CHD.

In general, this study was well conducted, and the manuscript was well written. I have the following comments.

(1) Abstract. Please describe the distribution of age and gender in both groups briefly in the Results.

RESPONSE: We appreciate the review for this comment. In the revised manuscript (Page 1, Line 56), we have now added the following sentence into the abstract:

"Of the 94 cases, the median age was 72.5 years old (IQR=16), and 59.6% were male, while in the control group the median age was 67 years old (IQR=22) and 64.6% were male."

(2) Table 1. There were three columns of HRs for multivariate model. Please describe what variable each column was adjusted for.

RESPONSE: In table 2 multivariate model, each column represents one model: we listed the HR of target predictor as well as all the adjusted covariates. For example, the 1st multivariate model used total number of comorbidities as the target predictor, adjusting for age, sex and time before 01/22/2020. To reduce the confusion, we have now added NA in those covariates that were not included in the model.

(3) Discussion. The authors did not recruit treatment into the analyses, which may confound the results. Please explain this issue in the Discussion.

RESPONSE: We appreciate the reviewer's comment. We acknowledge that, due to the lack of information of treatment in the health reports published by the local health commission websites, we did not include the treatment information into the model, which may produce some confounding effects. In the revised manuscript (Discussion section, Page 12, Line 429), we have acknowledged this as a limitation of the study, and pointed to the direction of future research:

"Moreover, due to the lack of information of treatment in the health reports published by the local health commission websites, we did not include treatment information into analysis, which may produce confounding effects. It calls for future research to investigate the mortality risk effect of pre-existing comorbidities by adding treatment as a covariate in the model."

Reviewer: 2

Reviewer Name: Priscila Maria Stolses Bergamo Francisco
Institution and Country:
State University of Campinas
São Paulo/Brazil

General comments

The study addresses a relevant question to Public Health today and assesses the relationship between pre-existing comorbidities and mortality from COVID-19, through a nested case-control design and public data from patients with confirmed SARS-CoV infection -2 in mainland China outside of Hubei Province. Secondary public domain data were used and the data collection, for each case, was properly described in stages. The model's assumptions were verified to perform statistical analyzes. The method is clearly described and the discussion is well structured.

In Introduction, the first paragraph, the authors could update the number of cases and countries affected around the world, as well as in the 2nd paragraph, including the reference (WHO).

RESPONSE: We appreciate this advice from the reiviewer. In the Introduction section, we have updated the information of cases and affected countries around the world, and have cited the most updated situation report of WHO (Page 3, Line 107):

"According to the COVID-19 situation reports of WHO, as of June 13th, 2020, the infection has caused 83,132 confirmed cases in mainland China, including 4,634 deaths. Internationally, a total of 7.41 million confirmed cases have been reported from 186 countries outside of China, including 418,000 deaths. Considering the global public health threat posed by COVID-19, unraveling the prognostic factors for patients, especially the risk factors of mortality associated with COVID-19, has important implications for clinical practice and is urgently warranted."

Specific comments

Model Results

(1) Paragraph 2: "Table 2 presents the results of univariate and multivariate weighted Cox models. Older age was associated with significantly higher death risk with similar magnitude in univariate and multivariate models. In the adjusted model, every 5-year increase in age was associated with na estimated 20% higher risk of death (p<0.001). No significant hazard difference was found between male and female patients. Disease infection during the early no-intervention period was associated with a higher risk of death, although not statistically significant".

I suggest: "Table 2 presents the results of univariate and multivariate weighted Cox models. Older age was associated with significantly higher death risk with similar magnitude in univariate and multivariate models. No significant hazard difference was found between male and female patients.

Although the disease infection during the early nointervention period was associated with a higher risk of death (pontual estimate), it was not statistically significant."

"In the adjusted model, every 5-year increase in age was associated with an estimated 20% higher risk of death" - **should not be related with Table 2, since this result is not represented there.**

RESPONSE: We thank the reviewer for these suggestions. We apologize for the confusion about the 5-year mortality risk of age. We interpreted the original 1-year mortality result from Table 2 to a 5-year scale (HR of age, 1.04 [95% CI, 1.02-1.06]; p<0.001). We have revised these sentences as follows:

"Table 2 presents the results of univariate and multivariate weighted Cox models. Older age was associated with significantly higher death risk with similar magnitude in univariate and multivariate models. In the adjusted model, every 1-year increase in age was associated with an estimated 4% higher risk of death (p<0.001). No significant hazard difference was found between male and female patients. Although the disease infection during the early nointervention period was associated with a higher risk of death (HR, 1.21 [95% CI, 0.74-1.98]; p=0.45), it was not statistically significant."

(2) Paragraph 3: "The increasing number of cumulative comorbities as associated with higher mortality risk in both unadjusted and adjusted models (p<0.001) **- In Table 2 and Table S2 the**

authors present the variable "Total numbers of comorbidities", but does not show that the

cumulative number of comorbities is related to higher mortality, since this variable does

not represent an increasing number of cumulative comorbities (gradient) - I suggest rewriting the sentence or showing these results in Table S2.

RESPONSE: In the revised manuscript, we have renamed the variable "total numbers of comorbidities" as comorbidity score. It was the summation of all nine comorbidities listed in Table 1 (CHD, hypertension, cardiac failure, cerebral infarction, chronic bronchitis, COPD, diabetes, renal failure and history of surgery), ranges from 0 to 9. We have revised this sentence as follows:

"In a separate model using comorbidiy score as predictor, we observed that higher comorbidity score was associated with higher mortality risk in both unadjusted and adjusted models (p<0.001 and p=0.001, respectively)."

(3) Paragraph 3: All preexisting comorbidities were associated with a higher risk of COVID-19 mortality in the univariate model, of which CHD had the largest hazard ratio (HR) of 4.2 (p<0.001), followed by cerebral infarction (HR=2.9, p=0.004), COPD (HR=2.6, p=0.01), renal

failure (HR=2.3, p=0.09), cardiac failure (HR=1.9, p=0.1), history of surgery (HR=1.7,

p=0.34), hypertension (HR=1.4,p=0.17), diabetes (HR=1.1, p=0.61) and chronic bronchitis (HR=1.1, p=0.55), **but not all statistically significant.**

RESPONSE: We appreciate the reviewer's advice. To be more specific and accurate, we have now revised this sentence as follows:

"All pre-existing comorbidities had hazard ratio (HR) over one in the univariate model, of which CHD had the largest HR of 4.2 (p<0.001), followed by cerebral infarction (HR=2.9, p=0.004), COPD (HR=2.6, p=0.01), renal failure (HR=2.3, p=0.09), cardiac failure (HR=1.9, p=0.1), history of surgery (HR=1.7, p=0.34), hypertension (HR=1.4, p=0.17), diabetes (HR=1.1, p=0.61) and chronic bronchitis (HR=1.1, p=0.55), but not all statistically significant."

Reviewer: 3
Reviewer Name: Moritz Tobiasch
Institution and Country: Landeskrankenhaus Hall in Tirol, Department of Medicine, Tyrolia, Austria

Thank you very much for the opportunity to review this paper. From a methodological point of view, it is quite obvious (and very welcome) that a fair amount of work and consultation with experts in biostatistics was put in this project.

As I was asked to primarily focus on statistical questions, I would like to specifically address this field - however, there is one more general objection that need to be clarified all along: The process of data gathering via news outlets and websites is all but unconventional for people accustomed to EU data protection standards. In my opinion, it would be advisable to clarify whether there was supervision by any sort of ethics committee.

RESPONSE: We appreciate the reviewer for this comment. We understand and highly value the importance of protecting patient privacy. The data published in news reports and websites were free of personally identifiable information and were open to the public. Our study was approved by Shanghai Jiao Tong University Public Health and Nursing Medical Research Ethics Committee (SJUPN-202001). News and public health reports were also used in previous research to investigate the epidemiology of COVID-19. Below is an example:

Lauer S. A., Grantz, K. H., Bi, Q., Forrest K. J., Zheng, Q., Meredith, H., Azman, A., Reich, N. G., & Lessler, J. (2020) The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine* 2020; 172: 577-582.

As far as statistics are concerned, I would like to ask, with all due respect, for some clarifications:
(1) In the methods section, it is mentioned that for continuous variables, a t-test was used. This test has gone a little out of fashion as it relies on assumptions that are only rarely met, namely, normal distribution and equal variances in both groups. Have all continuous variables been tested for normality, e.g., by Komolgorow-Smirnow or Shapiro-Wilk tests? If the baseline assumptions of normality and equal variances are not met, I would advise to use a non-parametric test.

RESPONSE: We appreciate the reviewer for this suggestion. Age was the only continuous variable included in the analysis. Before applying t-test, we checked QQ-plot (pasted below), the normality of which looked okay. But to be more conservative and more precise, we applied Mann–Whitney U test and achieve the same result (p<0.001), a non-parametric test to compared means as suggested by the reviewer. We also replaced "t-test" with "the non-parametric Mann–Whitney U test" in the text.
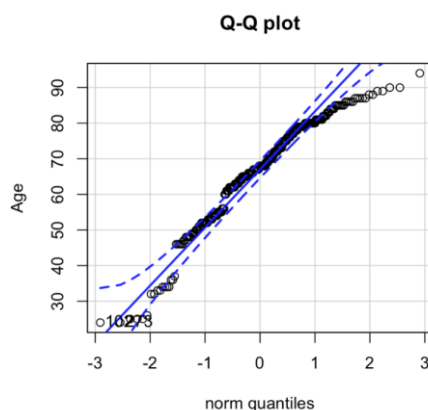
(2) I am quite impressed to see such an analysis performed in R. R is more and more used in a dedicated integrated development environment (IDE) such as RStudio. If this was the case, I would like to ask to state this as well. Furthermore, one package was explicitly mentioned and cited, which I greatly appreciate. If other packages were used as well, I would like to cite them as well.

RESPONSE: We thank the reviewer's comment. The analysis was indeed performed in R through RStudio version 1.2.5042. We have now included all the data and code (including R packages used) in the analysis online at Github:

https://github.com/GuTian-TianGu/COVID-19_NCCstudy

Moreover, in the revised manuscript (Page 7, Line 237), we have clarified the software and included the corresponding citation. The addd text reads:

"Analyses were performed in R 3.6.2 (R Foundation for Statistical Computing) through RStudio version 1.2.5042. Data and code are available online at GitHub."

(3) Is it possible to clarify how preexisting conditions were defined, e.g., for heart failure, was there a LVEF cutoff, or, for hepatic failure, was a classification according to AASLD or EASL or APASL used - or did all these data solely rely on the discretion of treating physicians or the press? The latter would not be a particular worrisome challenge, it just needs to be stated.

RESPONSE: We appreciate the reviewer for this comment. Since the publicly available case reports did not include diagnoses as detailed as in electronic health record, we were not able to classify comorbidities using laboratory test results. The data of patients' comorbidities solely relied on the information disclosed in the public health reports. We have now clarified this in the revised manuscript (Page 6, Line 208):

"Pre-existing comorbidities were recorded based on the description of case reports."

(4) Demographics: Can you comment on the fact why it happened that even after matching, the control cohort proved to be significantly younger than the case cohort?

RESPONSE: The significant difference of age between two groups was due to the imperfect matching on sex and age $\pm$ 1 year old instead of the exact matching. This was also the reason we still adjusted for age and sex in the model after matching. In addition, each control can be matched more than once in the nested case control (NCC) design (equivalent to sampling with replacement). As we noted in the manuscript (Page 5, Line 173): "NCC is cost-effective in data collection, and is especially suitable for research on the death risk of diseases such as COVID-19, where the number of event-free people largely exceeds those who experienced events"

(5) The stepwise risk stratification for age with increments of 5 years in a cohort with a median age of 70.7 is a little unprecise, in my view. In my country, the average age of death in males is currently 82 years, which makes up for only a little more than two incremental steps in your model, and I would think that even without COVID-19, between the ages of 70.7 and 82, the death rate might rise by 20% every 5 years.

RESPONSE: We modeled age as a continuous variable. As listed in Table 2, the hazard ratio of age indicates that with every one year increase in age, the estimated risk of COVID-19 associated mortality will increase 4%. Equivalently, we interpreted this result in a 5-year unit (HR per 5 years was exp(0.037268*5)=1.20) as every five years increase in age will lead to an estimated 20% risk of COVID-19 related death. Since the overall age in our study ranges from 24 to 94, we reported age per 5-year unit. To reduce confusion, we changed the age interpretation based on 1-year unit. For example (Page 9, Line 307):

"In the adjusted model, every 1-year increase in age was associated with an estimated 4% higher risk of death (p<0.001)."

(6) Your statement of equal risks of death between male and female patients is interesting and needs some discussion, as per usual assumption in the public (and hereby I mean specialists in critical care medicine treating COVID-19 patients) males are more often dying from COVID-19 than women. Is there, to your knowledge, data available reproducing your finding?

RESPONSE: We thank the reviewer for this point. In the revised manuscript (Discussion section, Page 11, Line 404), we have added a discussion of our results regarding non-significant effect of gender on mortality risk of COVID-19, and mentioned literature about gender difference in death risk associated with COVID-19. The added text reads:

"Previous studies yielded mixed results regarding gender differences in mortality risk of COVID-19. Some studies found male sex was associated with higher death risk of COVID-19 (e.g., Li et al., 2020), whereas other studies did not find gender to be a significant factor predicting the mortality risk of COVID-19 (e.g., Zhou et al., 2020; Ruan et al., 2020). In the current study, gender was not a significant mortality risk factor for COVID-19. It calls more future research to further our understanding of gender difference in the outcome of COVID-19 and the underlying mechanism."

(7) The last and possibly most important point is still a statistical one, but it happens to affect basically all studies on acute critical conditions. Your model of a 65 year old female with CHD and COVID-19 comes up with quite harsh prognostics, with a median (!) survival of mere 34 days. From my humble experience, I would rather say that this might be a selection or matching artifact (by the way, the "German anomaly" with exceptionally low mortality despite equal infection numbers in GER compared to other EU countries like ITA, ESP, FRA is possibly quite the inverse effect). May I ask to go through your matching process to find out why exactly such a short survival time was found? What is the time course and prognosis in oligosymptomatic COVID-19 patients with CHD?

RESPONSE: In the original manuscript, we provided the estimated median survival time (34 days) as well as the estimated 30-day survival probability (0.53 [95% CI, 0.34-0.82]) for a profile patient (65-year-old female with no other comorbidities) to illustrate the poor prognosis of patient with pre-existing CHD from the modeling results. The median survival time had a skewed distribution with an infinit 95% CI upper bound due to limited sample size. To reduce confusion, we decided to remove the estimated median survival day, and only report the equivalent 30-day survival probability with a normal 95% CI.

As we also mentioned in our reponse to comment (4), the control was sampled with replacement in the matching process in the Nested Case Control (NCC) design, which reflects as the weights. Thus, we listed the weight for each patient in the supplymentary material (https://github.com/GuTian-TianGu/COVID-19_NCCstudy). As described in method section in the text, "only survivors were assigned weights, since all cases (deaths) were included as designed with a weight of one"

(8) Additionally, I would like to see at least one typo addressed:
- p12 line 14: "plaque" instead of "plague"

RESPONSE: We appreciate the reviewer for pointing out this typo. We have corrected this typo in the text.

(9) As far as references are concerned, in the last days epidemiologic evidence from EU countries, UK, and USA emerged. For an even better discussion, I would like to ask whether the work presented here could be put in some context. Furthermore, a current hypothesis in COVID-19 induced critical illness is a generalised affection of endothelia, which makes preexisting vascular conditions a at least hypothetically dangerous influencing factor. Maybe the discussion could be expanded to these recent findings.

RESPONSE: We thank the reviewer for these insightful comments. In the revised manuscript (Discussion section, Page 11, Line 416), we have added a discussion of the contextual information of our findings:

In total, the work is very balanced, thorough, and interesting. After addressing the above-mentioned points, the paper might be ready for publication.

Reviewer: 4
Reviewer Name: Luis Puente-Maestu
Institution and Country: Hospital Universitario Gregorio Marañón-Facultad de Medicina de la Universidad Complutense de Madrid

(1) The study has some flaws. The most important is the large number of individuals with incomplete data and the important number of excluded cases for the final analysis, since it questions the validity of the study and it seems to this reviewer difficult to solve.

RESPONSE: We thank the reviewer for this comment. Due to the fact that COVID-19 has high infection rate yet low death rate, it is understandable that publicly reported cases do not cover all survivors, and thus we were not able to collect a large number of survivor information but mostly death cases. Therefore, we chose to use the nested case control (NCC) design, which perfectly fit this situation and could largely solve the potential bias issue.

NCC design is difference from the traditional case control, where each case is matched with one control on pre-specified criteria. As described in Introduction (Page 5, Line 173), "*NCC, also called risk set sampling, is prefered in the situation where the number of event-free people largely exceeds those who experienced events,*" such as our COVID-19 study where the prevalence of death was very low in the study population. In a typical NCC, all cases are included in the study while certain numbers of control are matched with replacement (i.e. each control can be matched to more than one case).

We have collect almost all the reported cases with patient-level information across the news outlets and official reports to the best of our ability. Due to the nature of NCC study, it's crutial to include all cases and prepare a representative number of controls as candidates to match. We were able to collect 84% (94/112) of the deaths in the time frame and 3% (354/12,861) of the survivors as candidates of control before matching. No evidence shown that the 18 missing deaths (9 from Heilongjiang, 5 from Henan, 3 from Beijing, and 2 from Hunan) were different from other deaths. Thus, we assumed they were missing at random, as well as the missing survivors. Since we were not able to dig further to confirm these assumptions, we listed it as a limitation in the Discussion (Page 12, Line 440):

"Lastly, we were not able to verify the missing at random assumption of the 18 missing deaths (9 from Heilongjiang, 5 from Henan, 3 from Beijing, and 2 from Hunan) as well as the survivors."

We have tried to collect the data as complete as we could, including the dynamic tracking method we introduced in Introduction. In other published researches that used similar data source, much more fewer cases were included in the modeling. For example, in the paper [1] below, among 780 deaths upon February 8th in Hubei Province, authors collected 48 deaths from National/Provincial health committee (same resource as ours), of which 13 were excluded due to missing time-covariates, 8 were exluded for other reasons, and only 24 left were used to to estimate the onset-to-death distribution.

In addition, as listed in Discussion, "*researchers have pointed out that severe cases may be over-represented in publicly reported data [2]*." Since pre-existing health coditions have been widely identified to be associated with severe cases [3-5], if the over-representaion of severe case existed, our results would reveal the correct direction but underestimate the true underlying risk relationship between pre-exisitng comorbidity and COVID-19 mortality, which hope to ease the reviewer's concern to some extent.

[1] R. Verity, L. C. Okell, I. Dorigatti et al., Estimates of the severity of coronavirus disease 2019: A modelbased analysis. *Lancet Infect. Dis.* 10.1016/S1473-3099(20)30243-7 (2020).

[2] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med* 2020.

[3] Mao R, Liang J, Shen J et al . Implications of COVID-19 for patients with pre-existing digestive diseases. *Lancet Gastroenterol.* Hepatol. 2020; 5: 426–428.

[4] Dietz W, Santos-Burgoa C. Obesity and its Implications for COVID-19 Mortality. *Obesity (Silver Spring).* 2020 Apr 1. doi: 10.1002/oby.22818.

[5] G.Targher et al. Patients with diabetes are at higher risk for severe illness from COVID-19. *Diabetes & Metablism.* 2020 May. https://doi.org/10.1016/j.diabet.2020.05.001

(2) Even in waivered from informed consent, the study should have been reviewed by an ethics in human research committee.

RESPONSE: We thank the reviewer for the comment. Our study was reviewed and approved by Shanghai Jiao Tong University Public Health and Nursing Medical Research Ethics Committee (SJUPN-202001).

(3) Page 6, line 32: "The data collection procedure was blinded to patient comorbidity information". Clarify; how the data collection could be blinded to comorbidities if they were considered

exposition factors

RESPONSE: We apologize for the confusion. We actuallyl meant that we collected all the available data regardless of patients' pre-existing comorbidity conditions. We have now clarified this in the revised manuscript (Page 5, Line 190) as follows:

"To avoid selection bias due to intentionally collecting patient with certain pre-existing comorbidities, two authors independently collected, compared and reviewed the full text of each case report."

(4) Page 6, line 40: was matched with up to three controls on gender and age ± 1 year old (94 cases, 181 controls). The authors should explain bit more in depth how the sample was collected, randomly among all the possible controls fulfilling the matching criteria, stratified by regions. Furthermore from figure 1 it seem that most of the available control population (i.e 12555/12861) had incomplete records. If that is true the validity (representativeness) of the study sample would be largely questionable

RESPONSE: We thank the reviewer for the comment. As described in the Statistical Analysis section, the matching procedure and the calculation of weights were simultaneously conducted "*via KMprob function in multipleNCC R package, by specifying the Kaplan-Meier type weights with additional matching on gender and age ± 1 year old. Only survivors were assigned weights, since all cases (deaths) were included as designed with a weight of one.*" As we have already mentioned in our response to comment (1), in a typical NCC, all cases are included in the study while certain numbers of control are matched with replacement (i.e. equivalent to sampling with replacement). We initially included the explaination of NCC design, and removed later after seeing several NCC papers published on BMJ Open without explaining. To reduce the possible confusion, we added the following explaination in the Method in the text (Page 5, Line 192):

"Following the typical NCC design setting where all events are included and a certain numbers of control are matched with replacement, all deaths were included as cases, and each case was matched with up to three controls on gender and age $\pm$ 1 year old (94 cases and 181 controls)"

As the reviewer pointed out in Figure 1, among all the 12,973 infected patients across 31 province-level regions in mainland China (excluding Hubei Province) upon March 8th, we have collect almost all the reported cases with patient-level information across the news outlets and official reports to the best of our ability. As we have already mentioned in our response to comment (1), due to the nature of NCC study, it's crutial to include all cases and prepare a representative number of controls as candidates to match. Although not able to verify, we assumed the missing survivors were missing at random, and listed as a limitation in Discussion.

(5) Page 6, line 46: We routinely searched for daily news and public health reports on confirmed COVID-19. Are the final sources of information available for other investigators to check your results?

RESPONSE: We have uploaded all the patient-level data, code to replicate the analysis results, and all the data sources with specific website links on Github (https://github.com/GuTian-TianGu/COVID-19_NCCstudy). Those information assures that other investigators can reproduce our results.

(6) Page 7, line 8: "who had complete information on basic demographics (age, gender and region), disease onset date--the first time a patient became symptomatic, and history of comorbidities (include but not limited to hypertension, cardiovascular disease, diabetes and respiratory diseases) were included in the analysis" Was the proportion of excluded subjects even among regions, or said in other words, had the registry the same quality among regions?. Had all the included individuals the same chance of being diagnosed of some of the comorbidities with independence that the cases came for urban and rural areas, more or less developed zones?

RESPONSE: We appreciate the reviewer's insightful comment. As listed in Appendix Table A1, from each province, we included around 0.5%-3% of the total provincial survivors, except Shangdong Province who contributed 10% (66/692), mainly due to detailed reports available from Shandong Province. In an additional sensitivity analysis (result not shown here), we included Shandong as an indicator in all models. The significance of all covariates remained the same, and the HR disparity was acceptable.

Regarding potential differences between rural and urban areas, we were not able to distinguish exactly urban and rural areas from the data. But to address this concern, in an additional sensitivity analysis, we adjusted for the west area, where socioeconomic development (e.g., annual income and economic growth) were relatively backward compared with the rest of the China. According to National Statistics Bureau of China, we defined west area to include the following provinces: Neimenggu, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Xizang, Shanxi, Gansu, Qinghai, Ningxia, and Xinjiang. There were a total of 20 (7.3%) survivors and 21 (22.3%) deaths came from the west area in the dataset. The sensitivity model results listed below were consistent with our models in the main text:

Characteristic	Multivariate Weighted Cox Regression					
	HR (95% CI)	P Value	HR (95% CI)	P Value	HR (95% CI)	P Value
Age	1.04 (1.02-1.07)	**<0.001**	1.04 (1.02-1.06)	**<0.001**	1.04 (1.01-1.06)	**<0.001**
Male	1.07 (0.68-1.69)	0.49	1.09 (0.70-1.71)	0.70	1.00 (0.63-1.60)	0.99
Before 01/22/2020	1.25 (0.77-2.03)	0.28	1.13 (0.69-1.85)	0.63	1.15 (0.70-1.89)	0.45
West area	1.51 (0.86-2.64)	0.15	1.77 (1.06-2.97)	**0.03**	1.69 (1.01-2.83)	
Comorbidity score	1.29 (1.10-1.52)	**0.002**	NA	NA	NA	NA

CHD	NA	NA	3.12 (1.88-5.16)	**<0.001**	3.11 (1.88-5.12)	**<0.001**
Cerebral infarction	NA	NA	NA	NA	1.95 (0.99-3.86)	0.054
COPD	NA	NA	NA	NA	1.73 (0.81-3.69)	0.16
Renal failure	NA	NA	NA	NA	1.87 (0.75-4.69)	0.18

HR=hazard ratio; CHD=coronary heart disease; COPD=chronic obstructive pulmonary disease.

Bold: statistically significant using threshold p<0.05.

(7) Page 7, line 13 "include but not limited to hypertension, cardiovascular disease, diabetes and respiratory diseases) were included in the analysis" all the comorbidities included in the analysis should be shown here or in the supplementary material. AS per table 2 they are not many more.

RESPONSE: We thank the reviewer for the comment. We have listed all the comobidity categories used in the analysis in Table 1, including hypertension, CHD, cardiac failure, cerebral infarction, diabetes, chronic bronchitis, COPD, renal failure and hepatic failure. In the revised manuscript (Page 6, Line 211), we have now clarified this sentence as follows:

"History of comorbidities (include hypertension, CHD, cardiac failure, cerebral infarction, diabetes, chronic bronchitis, COPD, renal failure and hepatic failure)."

(8) Page 7, line 13: Asymptomatic patients were not included in this study. This must be moved to the study design section. A reference to this criteria has to be included in the abstract and somehow if possible in the title.

RESPONSE: We appreciate the reviewer's advice. In the revised manuscript, we have now added the following sentence to the abstract (Page 2, Line 48):

"Asymptomatic patients were not considered in this study."

(9) Page 8, line 6 The matching between cases and controls, and relative weights were simultaneously obtained via KMprob function in multipleNCC R package" The time-fixed IP exposure weights should be presented somewhere

RESPONSE: We thank the reviewer for the suggestion. We have now listed the weight for each patient online in Github (https://github.com/GuTian-TianGu/COVID-19_NCCstudy), along with all the data and codes we used in this study.

(10) Page 8, line 10: Asymptomatic patients were not included in this study

RESPONSE: As we have also mentioned in our response to comment#8, we have added the following sentence to the abstract "*Asymptomatic patients were not considered in this study.*"

(11) Page 8, line 12: Those survivors with sampling probabilities of zero were considered as "fail to match" and excluded from the study. The number of excluded patients and the region they come from should be shown.

RESPONSE: We thank the reviewer for the comment. We listed the weight for each patient online in Github (https://github.com/GuTian-TianGu/COVID-19_NCCstudy), including those who had zero weights. Moreover, we have changed the (Page 7, Line 248):

"A total of 113 survivors (mean age 46.5) with sampling probabilities of zero were considered as 'fail to match' and excluded from the study, mainly due to younger age than cases. A majority of the excluded patients were from Shandong Province (38.1%) due to relatively high representation of the sample (detailed information of excluded survivors is available online at Github). In a sensitivity analysis adjusting for Shandong Province (results not shown here), we observed the consistent results as the main analysis."

(12) Page 8, line 15 The total number of comorbidities was defined as the summation of comorbidities, ranging from zero to four or above." Was this resulting variable coded for multivariable analysis (i.e $X_i=1$, the rest 0, resulting in k-1 coded variables). Besides it seems from your results (see below), that the number of comorbidities is not a homogeneous risk factor. You may approach to comorbidity analysis by calculating the Carlson index which tries to account for the difference impact of different comorbidities in survival.

RESPONSE: We thank the reviewer for the comment. We calculated "the total number of comorbidities" as the summation of all 9 comorbidities listed in Table 1 (CHD, hypertension, cardiac failure, cerebral infarction, chronic bronchitis, COPD, diabetes, renal failure and history of surgery). We renamed it as "comorbidity score", which ranges from 0 to 9. Instead of listing counts (%), we considered it as a continuous variable in Table 1.

Since the publicly available case reports did not include diagnoses as detailed as in electronic health records, we were not able to obtain information such as diabetic end organ damage or metastatic status of tumor, which were the essential elements to calculate Charlson Comorbidity Index (CCI). Thus, we compromised to calculate a simple composite comorbidity score instead of a more comprehensive one like CCI.

(13) Table 2 Some HR are shown as 1.0 (1.0-1.1), please give the number of decimals needed to show that both the HR and the confidence interval do not include 1)

RESPONSE: We thank the reviewer for this suggestion. We have changed corresponding values to 2 or 3 decimals as needed in Table 2:

Table 2. Univariate and multivariate model result from weighted Cox proportional hazard regression

Characteristic	Univariate				Multivariate			
	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P Value	HR (95% CI)	P
Age	1.05 (1.0-1.1)	**<0.001**	1.05 (1.03-1.07)	**<0.001**	1.0 (1.0-1.1)	**<0.001**	1.04 (1.02-1.06)	**<0.001**
Male	0.76 (0.5-1.2)	0.24	1.17 (0.74-1.85)	0.49	1.1 (0.7-1.7)	0.71	0.997 (0.62-1.60)	0.99
Before 01/22/2020	1.12 (0.7-1.8)	0.66	1.31 (0.80-2.13)	0.28	1.2 (0.7-1.9)	0.48	1.21 (0.74-1.98)	0.45
Comorbidity score	1.50 (1.27-1.75)	**<0.001**	1.31 (1.11-1.54)	**0.001**	NA	NA	NA	NA
Cardiocerebrovascular								
CHD	4.19 (2.5-7.1)	**<0.001**	NA	NA	2.9 (1.7-4.9)	**<0.001**	3.01 (1.82-4.98)	**<0.001**
Hypertension	1.37 (0.9-2.2)	0.17	NA	NA	NA	NA	NA	NA
Cardiac failure	1.85 (0.9-3.9)	0.10	NA	NA	NA	NA	NA	NA
Cerebral infarction	2.86 (1.4-5.8)	**0.004**	NA	NA	NA	NA	1.90 (0.94-3.8)	0.07
Respiratory								
Chronic bronchitis	1.05 (0.4-2.5)	0.55	NA	NA	NA	NA	NA	NA
COPD	2.61 (1.2-5.6)	**0.01**	NA	NA	NA	NA	1.85 (0.89-3.85)	0.10
Endocrine								
Diabetes	1.14 (0.7-1.9)	0.61	NA	NA	NA	NA	NA	NA
Others								
Renal failure	2.30 (0.9-6.0)	0.09	NA	NA	NA	NA	2.02 (0.81-5.07)	0.13
History of surgery	1.71 (0.6-5.1)	0.34	NA	NA	NA	NA	NA	NA

HR=hazard ratio; CHD=coronary heart disease; COPD=chronic obstructive pulmonary disease.

(14) Pag 11, line 22 The increasing number of cumulative comorbidities was associated with higher mortality risk in both unadjusted and adjusted models (p<0.001). This statement may not be true. According to your data does not seem that it is the same to have two comorbidities that were not CHD, that only one which was CHD.

RESPONSE: We appreciate the reviewer's comment. The statement that "The increasing number of cumulative comorbidities was associated with higher mortality risk in both unadjusted and adjusted models (p<0.001)" was based on the result of the first multivariate model in Table 2 (predictors: comorbidity score + age + gender + before 01/22/2020, attached in response of comment 13). In this revision, we modeled it as a continuous variable to measure patients' overall health condition, instead of using ordinal variable in the original manuscript. We agree with the reviewer that the mortality risk may not increase linearly as the number of comorbidity increases. Thus, instead of interpreting it as how much mortality risk increase when patient had an additional comorbidity, we have revised this statement as follows (Page 9, Line 311):

"In a separate model using comorbidiy score as predictor, we observed that higher comorbidity score was associated with higher mortality risk in both unadjusted and adjusted models (p<0.001 and p=0.001, respectively)."

(15) Page 11, line 41 "All preexisting comorbidities were associated with a higher risk of COVID-19 mortality in the univariate model," As far as table 2 reports this is not true. Some commodities were not found significantly associated to survival expectancy

RESPONSE: We thank the reviewer for pointing this out. In Table 2 listed above, all univariate model had estimated HR >1, although some were not statistically significant. To be more specific and accurate, we have now revised this sentence as follows:

All pre-existing comorbidities had hazard ratio (HR) over one in the univariate model, of which CHD had the largest HR of 4.2 (p<0.001), followed by cerebral infarction (HR=2.9, p=0.004), COPD (HR=2.6, p=0.01), renal failure (HR=2.3, p=0.09), cardiac failure (HR=1.9, p=0.1), history of surgery (HR=1.7, p=0.34), hypertension (HR=1.4, p=0.17), diabetes (HR=1.1, p=0.61) and chronic bronchitis (HR=1.1, p=0.55), but not all statistically significant.

VERSION 2 – REVIEW

REVIEWER	Moritz Tobiasch LKH Hall in Tirol
REVIEW RETURNED	12-Jul-2020

GENERAL COMMENTS	The authors thankfully have addressed the majority of the issues raised by the reviewers in the first round. As I was specifically asked

	to review the statistical workup, I would like to point out that reporting depth and quality has indeed been markedly improved with a full disclosure of the R markdown document. This should be the standard of how statistical computations are reported, I clearly approve of this effort. In my view, two central points still need clarification: if, while conducting a nested case-control study, age was one of the matching parameters, then how can it be explained that groups significantly differ in this point? Thankfully, a calculation corrected for age is provided, but I would like to see this effect discussed in the "limitations" section in the discussion. Second, ethical oversight of this project is not yet clear to me. If data were fully and entirely taken from publicly accessible, open databases, and this fully in agreement with chinese law, this should explicitly been stated, including a statement why no formal consultation of an ethics commitee was attempted. It still remains an issue when reporting these data, and efforts must be made to minimize any sort of intransparency in this regard. During the workover, some grammatical and expression errors were introduced, needing revision.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 3
Reviewer Name: Moritz Tobiasch
Institution and Country: LKH Hall in Tirol

Dear madam or sir,
the authors thankfully have addressed the majority of the issues raised by the reviewers in the first round. As I was specifically asked to review the statistical workup, I would like to point out that reporting depth and quality has indeed been markedly improved with a full disclosure of the R markdown document. This should be the standard of how statistical computations are reported, I clearly approve of this effort.

In my view, two central points still need clarification:
(1) if, while conducting a nested case-control study, age was one of the matching parameters, then how can it be explained that groups significantly differ in this point? Thankfully, a calculation corrected for age is provided, but I would like to see this effect discussed in the "limitations" section in the discussion.

RESPONSE: We appreciate the reviewer's comment. The significant difference of age between two groups was due to the imperfect matching on sex and age 1 year old instead of the exact matching (each control can be matched more than once in thenested case control design [NCC], which is equivalent to sampling with replacement). As pointed out by the reviewer, we adjusted for age and sex in all models after matching. In the revised manuscript, we have acknowledged this as a limitation of the study (Page 11 Line 355):

Following the NCC design, we allowed the controls to be matched with cases on age 1 year old instead of the exact matching, which caused the significant age difference between two groups (Table 1). Thus, we adjusted for the matching covariates in all the models to address this [32].

[32] Stoer, N. and Samuelsen, S. Inverse probability weighting in nested case-control studies with additional matching - a simulation study. *Stat Med* 2013;32, 5328-5339.

(2) Second, ethical oversight of this project is not yet clear to me. If data were fully and entirely taken from publicly accessible, open databases, and this fully in agreement with chinese law, this should explicitly been stated, including a statement why no formal consultation of an ethics commitee was

attempted. It still remains an issue when reporting these data, and efforts must be made to minimize any sort of intransparency in this regard.

RESPONSE: We appreciate the reviewer's comment. The Shanghai Jiao Tong University Public Health and Nursing Medical Research Ethics Committee reviewed and approved our study protocol (protocol number: SJUPN-202001) before the study was carried out. Since the data published on the national/provincial/municipal health commission websites were free of any identifiable personal information, our study involves no more than minimal risk, and does not adversely affect the rights and welfare of the participants. We have submitted the ethic approval letter for editor review.

In fact, the same data collection method (i.e., collecting data through reports published on the national/provincial/municipal health commission websites) was adopted by a few published studies on the epidemiology of the COVID-19 outbreak. We have provided a few example articles below:

- Pan A, Liu L, Wang C, et al. Association of Public Health Interventions with the Epidemiology of the COVID-19 Outbreak in Wuhan, China [published online ahead of print, 2020 Apr 10]. *JAMA*. 2020;323(19):1-9. doi:10.1001/jama.2020.6130

- Lauer S. A., Grantz, K. H., Bi, Q., Forrest K. J., Zheng, Q., Meredith, H., Azman, A., Reich, N. G., & Lessler, J. (2020) The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine 2020*; 172: 577-582.

- Roosa K, Lee Y, Luo R, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect Dis Model*. 2020;5:256-263. Published 2020 Feb 14. doi:10.1016/j.idm.2020.02.002

(3) During the workover, some grammatical and expression errors were introduced, needing revision.

RESPONSE: We thank the reviewer for the comment. We have carefully corrected the grammatical and language errors throughout the revised manuscript.